# Random and targeted transgene insertion in *Caenorhabditis elegans* using a modified *Mos1* transposon

Christian Frøkjær-Jensen[1–3], M Wayne Davis[1,2], Mihail Sarov[4], Jon Taylor[5], Stephane Flibotte[5], Matthew LaBella[2], Andrei Pozniakovsky[4], Donald G Moerman[5] & Erik M Jorgensen[1,2]

**We have generated a recombinant *Mos1* transposon that can insert up to 45-kb transgenes into the *Caenorhabditis elegans* genome. The minimal *Mos1* transposon (*miniMos*) is 550 bp long and inserts DNA into the genome at high frequency (~60% of injected animals). Genetic and antibiotic markers can be used for selection, and the transposon is active in *C. elegans* isolates and *Caenorhabditis briggsae*. We used the *miniMos* transposon to generate six universal *Mos1*-mediated single-copy insertion (mosSCI) landing sites that allow targeted transgene insertion with a single targeting vector into permissive expression sites on all autosomes. We also generated two collections of strains: a set of bright fluorescent insertions that are useful as dominant, genetic balancers and a set of *lacO* insertions to track genome position.**



Some DNA transposons can carry nontransposon DNA and still retain the ability to insert themselves randomly into chromosomal DNA. For example, the P element is used extensively to insert transgenes into the fruit fly *Drosophila melanogaster*[1]. The P element has also been used in the fly to generate large-scale gene knockout libraries, to drive tissue-specific expression using the Gal4 enhancer trap, to study genomic position effects and to generate targeted transgene insertion sites[2–5]. Similarly, other DNA-based transposons (such as *Sleeping Beauty*, *piggyBac* and *Tol2*) have successfully been used for transgenesis in a variety of genetically tractable systems including human tissue culture cells, mice, zebrafish, frogs and flies[6].

In *C. elegans*, transgenic animals are most frequently generated by DNA injection into the syncytial germ line to generate extrachromosomal arrays[7]. Biolistic transformation can be used for stable, but random, genomic integration of a single or a small number of plasmids[8]. The fly transposon *Mos1* is active in *C. elegans* but has limited cargo capacity (~500 bp) and is therefore not used directly for transgenesis[9]. Instead, excisions of *Mos1* inserts are used to generate double-strand DNA breaks, which are repaired from injected template DNA[10]. Through the use of positive and negative selection markers, a single copy of a transgene can be inserted into the genome directly via injection of mosSCI[11,12]. An alternative method to modify genomes that does not rely on transposons but on the bacterial clustered, regularly interspaced, short palindromic repeats (CRISPR)-Cas9 system[13] has recently been adapted for *C. elegans* to allow genome editing at endogenous loci[14–16].

Here we demonstrate that a modified *Mos1* transposon *miniMos* can carry large fragments of DNA, even 45-kb fosmids, into the *C. elegans* genome. We show that insertions can be selected using either genetic or antibiotic markers and that the transposon can be mobilized in wild isolates of *C. elegans* and *C. briggsae*. We have used *miniMos* to generate a set of strains with fluorescent markers that can be used as genetic balancers and *lacO* insertions that can track genome position in the nucleus. Furthermore, we have used the *miniMos* transposon to generate six universal mosSCI landing sites that allow insertion of a single transgene construct into permissive sites on all autosomes.

## RESULTS

### A recombinant *Mos1* element transposes with exogenous DNA

The requirements for transposition of mariner elements (*Mos1* and the closely related Peach transposon) vary depending on whether the transposon is embedded in chromatin or is contained within injected plasmid DNA. Mariner transposons within chromosomes require internal sequences to transpose[17] and can carry cargo only if the cargo is flanked by intact transposons[18]. By contrast, transposons injected as plasmids can transpose efficiently even if they contain internal deletions and carry cargo[19]. Experiments *in vitro* have further demonstrated that modifications to the inverted terminal repeats improve transposition frequency[20]. We tested whether modified *Mos1* elements and plasmid injection protocols[11] could overcome previously described limitations for *Mos1* transposition in *C. elegans*[9]. We generated a composite

[1]Howard Hughes Medical Institute, University of Utah, Salt Lake City, Utah, USA. [2]Department of Biology, University of Utah, Salt Lake City, Utah, USA. [3]Danish National Research Foundation Centre for Cardiac Arrhythmia, University of Copenhagen, Copenhagen, Denmark. [4]TransgeneOmics, Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany. [5]Department of Zoology, University of British Columbia, Vancouver, British Columbia, Canada. Correspondence should be addressed to E.M.J. (jorgensen@biology.utah.edu) or C.F.-J. (christianfj@gmail.com).

RECEIVED 4 DECEMBER 2013; ACCEPTED 29 JANUARY 2014; PUBLISHED ONLINE 16 MARCH 2014; DOI:10.1038/NMETH.2889

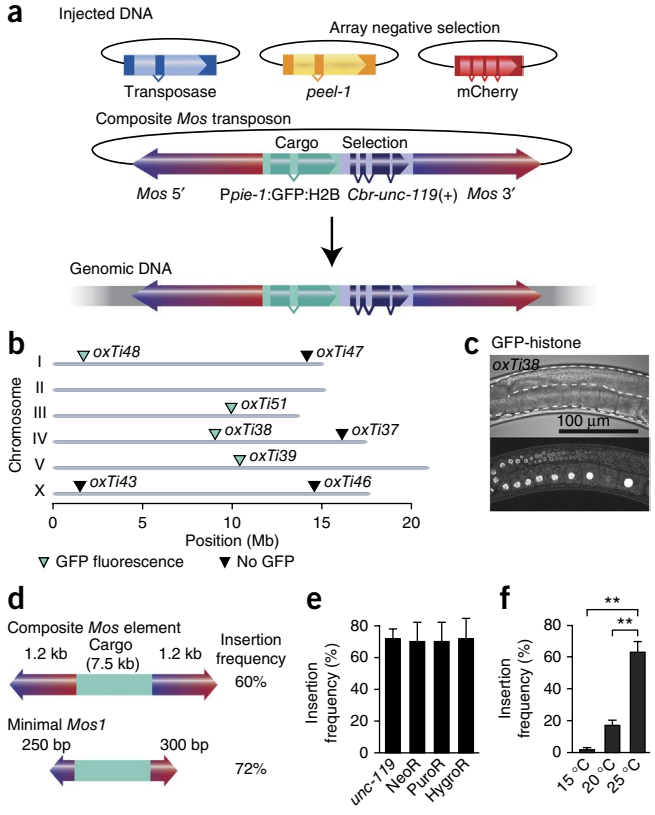


**Figure 1** | A modified *Mos1* transposon can carry cargo. (**a**) Schematic of the recombinant *Mos1* insertion protocol. Transposon DNA is co-injected with a helper plasmid expressing the transposase (P*eft-3*:*Mos1* transposase). Negative selection markers (P*hsp-16.41*:*peel-1*, P*myo-2*:mCherry, P*rab-3*:mCherry and P*myo-3*:mCherry) were used to select against array-bearing transgenic animals. (**b**) Genomic locations of insertions identified by *Cbr-unc-119*(+) rescue of *unc-119* mutants. All insertions rescued *unc-119*, but not all strains expressed GFP-histone in the germ line. Germline fluorescence is indicated with turquoise (GFP positive) or black (no fluorescence) triangles. (**c**) Fluorescence image of germline expression. Transposon insertion *oxTi38* expressed GFP-histone in the germ line (P*pie-1*:GFP:H2B). Top, differential interference contrast; bottom, confocal fluorescence image. (**d**) Schematic of the minimal *Mos1* transposon (*miniMos*). 550 bp was enough to retain full insertion frequency. (**e**) Insertion frequencies with the genetic marker *unc-119*(+) and antibiotic selection markers G418 (NeoR), puromycin (PuroR) or hygromycin B (HygroR). Each antibiotic was tested on animals injected on two different days. Values show the average from all injections ($n$ = 45–122 animals), and error bars show the 95% confidence interval (modified Wald method). (**f**) Insertion frequencies at different temperatures. Values shown are averages of three independent replicates (injections), and error bars represent s.e.m. Statistics: repeated measures ANOVA ($P$ = 0.0017) with Bonferroni *post hoc* comparison; $^{**}P < 0.01$.

*Mos1* transposon with a 7.5 kb transgene (containing P*pie-1*:GFP:histone and *Cbr-unc-119*(+)) and tested transposition by plasmid injection (**Fig. 1a** and **Supplementary Fig. 1**). We co-injected the composite *Mos1* transposon with a helper plasmid expressing the transposase and fluorescent extrachromosomal array markers. We injected 27 *unc-119* animals and identified 17 independent lines with recombinant *Mos1* insertions (62% $P_0$ insertion frequency). 47% (8 of 17) of the strains expressed GFP in the germ line (**Fig. 1c**). We mapped four GFP expressors and four non-expressors by inverse PCR[21] to unique insertion sites. Nonfluorescent insertions were found on autosomal arms, which have high levels of repressive chromatin marks[22], or the X chromosome, which is inactivated in the germ line[23] (**Fig. 1**). It is likely that these P*pie-1*:GFP:histone insertions are silenced through a combination of small RNAs that detect foreign DNAs and protect endogenous gene expression in the germ line[24–26] and subsequent modifications to the chromatin environment. We are currently characterizing germline and somatic position effects in detail (C.F.-J. and E.M.J., unpublished data).

The composite *Mos1* element was flanked by two essentially full-length *Mos1* elements. To identify a *miniMos* we tested transposition of truncated composite elements. Only 250–300 bp on either side was required for transposition with comparable efficiency to that of the composite transposon (**Supplementary Fig. 1**).

The composite transposon could also be mobilized from extrachromosomal arrays containing the transposon and the transposase under the control of a heat-shock promoter. From one extrachromosomal line (EG6346) we isolated two insertions from 300 heat-shocked animals (0.7%), and from a second line (EG6347) we isolated 12 insertions from 410 heat-shocked animals (2.9%). All insertions generated by mobilization from arrays were independent and mapped to unique genomic locations. It might be possible to generate large-scale transposon collections using a heat-shock protocol that are similar to the genome-wide collection of wild-type *Mos1* inserts[27]. However, it is currently more efficient to generate insertions directly by plasmid injection.

To determine whether composite *Mos1* insertions can be remobilized from genomic locations, we tried to remobilize the *oxTi51* insert by injection of the transposase gene and use of selection markers to detect germline excision and repair (**Supplementary Note**). We were unable to detect remobilization from 48 injections.

Thus, in agreement with experiments in flies[18–20,28]: (i) composite *Mos1* elements were able to transpose at high efficiency from injected plasmids and did not require most internal *Mos1* sequences, (ii) composite *Mos1* elements transposed at lower efficiency from extrachromosomal arrays and (iii) genomic insertions were not easily remobilized.

### Insertion into natural isolates and *C. briggsae*

We tested other genetic and antibiotic constructs as selectable markers for *miniMos* insertion. We generated insertions of otherwise identical constructs using *unc-119*(+)[29], G418 (NeoR)[30], puromycin (PuroR)[31] and hygromycin B (HygroR)[32] selection at similar frequencies (**Fig. 1e**). The genetic marker *unc-18*(+) was also as efficient as *unc-119*(+) selection (*unc-18*(+), 38%, $n$ = 13; *unc-119*(+), 34%, $n$ = 32) for a different construct. We were unable to generate insertions with two temperature-sensitive selection markers, *lin-5* and *spd-1*, that are necessary in the germ line. Insertions were probably not recovered because *miniMos* transposition was strongly temperature sensitive, with insertions occurring only at low frequency at 15 °C but at high frequency at 25 °C (2% at 15 °C, $n$ = 114; 62% at 25 °C, $n$ = 102) (**Fig. 1f**). Extrachromosomal arrays are generally silenced in the germ line[33], and injected DNA therefore cannot rescue *lin-5* and *spd-1* animals at 25 °C. Excision of the native *Mos1* element for mosSCI transgenesis at *ttTi5605* showed no temperature

**Table 1** | Recombinant *Mos1* transposon inserts at high frequency

| Injected $P_0$ animal no. | 1 | 2 | 3 | 4 | 5 | Total |
|---|---|---|---|---|---|---|
| Singled $F_1$ animals (rescued) | 24 | 45 | 40 | 18 | 29 | 156 |
| Insertions from rescued $F_1$ animals | 5 | 5 | 1 | 1 | 6 | 18 |
| Insertions from nonrescued $F_1$ animals | 0 | 1 | 0 | 0 | 1 | 2 |
| Single fluorophore | 5 | 6 | 1 | 1 | 7 | 20 |
| Multiple fluorophores | 0 | 0 | 0 | 0 | 0 | 0 |
| **Fluorescence of insertions** | | | | | | |
| P*eft-3*:GFP:H2B | 1 | 1 | 1 | 0 | 2 | 5 |
| P*eft-3*:mCherry | 2 | 3 | 0 | 1 | 2 | 8 |
| P*eft-3*:tdTomato:H2B | 2 | 2 | 0 | 0 | 3 | 7 |

Five *unc-119* animals were injected with a mix containing three *miniMos* elements carrying *Cbr-unc-119* and either P*eft-3*:GFP:H2B, P*eft-3*:mCherry or P*eft-3*:tdTomato:H2B transgenes. Three days later, a single $F_1$-rescued animal was picked to a new plate. One week later, plates were heat shocked to express PEEL-1 and kill array-bearing animals, and insertions from rescued $F_1$ animals were screened for the presence of single ("single fluorophore") or multiple ("multiple fluorophores") transgenes. All seven insertions from strain no. 5 mapped to independent genomic locations.

dependence (15% at 15 °C, $n$ = 71; 13% at 20 °C, $n$ = 75; 15% at 25 °C, $n$ = 71). It may be possible to use temperature-sensitive genetic markers such as *lin-5* or *spd-1* by injecting DNA into balanced strains that can be maintained at 25 °C.

We tested the $P_0$ insertion frequency into three highly diverged natural *C. elegans* isolates with NeoR selection: CB4856 (Hawaii), ED3040 (South Africa) and JU345 (France)[34]. The *miniMos* element was active in all strains although with variable insertion frequencies (6%, CB4856, $n$ = 17; 68%, ED3040, $n$ = 22; 16%, JU345, $n$ = 19). This variation might be due to differences in genetic backgrounds or differences in susceptibility to antibiotics[30]. *miniMos* could also be mobilized in other species. We successfully inserted a P*pie-1*:GFP:histone construct into a *C. briggsae* strain (6%, $n$ = 90) that was mutant for *Cbr-unc-119* (ref. 35); two of five animals showed stable GFP expression in the germ line. In an attempt to improve transposition efficiency in *C. briggsae*, we generated *cbr*-P*eft-3*:*Mos1* transposase and *cbr*-P*pie-1*:*Mos1* transposase constructs; however, the insertion frequency did not improve with either construct (0%, *cbr*-P*eft-3*, $n$ = 137 and 5%, *cbr*-P*pie-1*, $n$ = 43).

## Each strain contains a single *miniMos* insertion

To determine the insertion frequency in $F_1$ animals and the transgene copy number in each strain, we injected a mix of three different *miniMos* elements that could be distinguished by color (red or green) and cellular localization (cytosolic or nuclear) (**Table 1**). We injected five $P_0$ animals, picked 156 *unc-119* rescued $F_1$ animals to individual plates and recovered 20 independent insertions (11.5% $F_1$ insertion frequency). This frequency is comparable to the frequency of generating semistable transgenic animals by simple array injection (10%)[7]. All 20 insertions were fluorescent and expressed only one of the fluorophores from the injection mix (**Table 1**). Insertions from the same injected animal were independent; we determined all seven insertion sites from animal no. 5 by inverse PCR and all mapped to unique positions in the genome (*oxTi306*–*oxTi312*; **Supplementary Table 1**).

We also confirmed that insertion strains contain a single insertion by segregation in crosses (**Supplementary Note**). How can a single injection generate several independent insertions and yet each strain contain only a single insertion? We determined that this is possible because insertions were generated at relatively low frequency but occurred in the $F_1$ generation when the population expanded (**Supplementary Fig. 2**).

To facilitate identification of transposon insertion sites, we added new symmetric restriction sites to the *miniMos* vectors and optimized the inverse PCR protocols (**Supplementary Fig. 3** and **Supplementary Protocol**). We tested the optimized protocol in individual reactions and 96-well reactions on a collection of bright fluorescent P*eft-3*:tdTomato:H2B inserts (where tdTomato is tandem dimer Tomato and H2B is histone H2B), which will be useful as dominant chromosome balancers for *C. elegans* crosses (**Supplementary Fig. 4**).

12% of the inverse PCR reactions contained sequences from the injected plasmid backbone, a result indicating that some transpositions included two adjacent *miniMos* elements ('composite transposition'; **Supplementary Fig. 1**). Sequencing showed that the entire backbone of the injected plasmid had inserted. Incorporating the negative *peel-1* selection marker[11], which is heat-shock inducible, into the backbone of injected *miniMos* plasmids effectively selected against these types of complex insertions.

P-element transgenesis has been used to generate loss-of-function mutants in *Drosophila*[3]. Although we did not directly

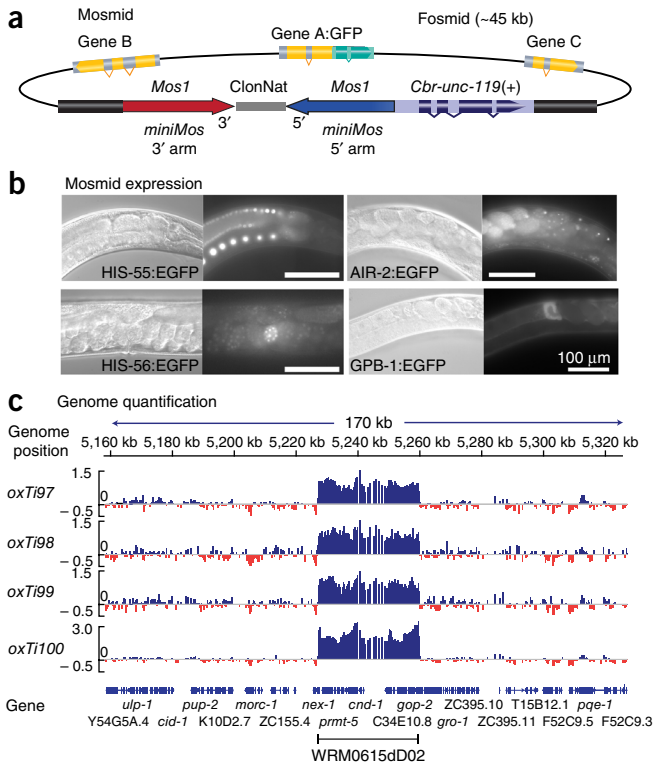


**Figure 2** | Fosmid insertions are intact. (**a**) Schematic of *Mos1*-based fosmids (Mosmids). *Mos1* and *Cbr-unc-119*(+) selection recombineered into the backbone of a fosmid carrying a GFP-tagged gene. (**b**) Fluorescence microscopy of Mosmid insertions. Four different Mosmid insertions with GFP show expression from the tagged genes. (**c**) Comparative genome hybridization (CGH) of genomic DNA from four independent insertions of the Mosmid WRM0615dD02 containing tagged *cnd-1*. CGH is based on dense oligonucleotide arrays tiled against a genome of interest and labeling of sample DNA and control DNA with different fluorophores. Genomic regions that differ between sample and control will show a difference in the ratio between the two color intensities. The Mosmid with *cnd-1*:EGFP contained an error rendering the fusion protein nonfluorescent.

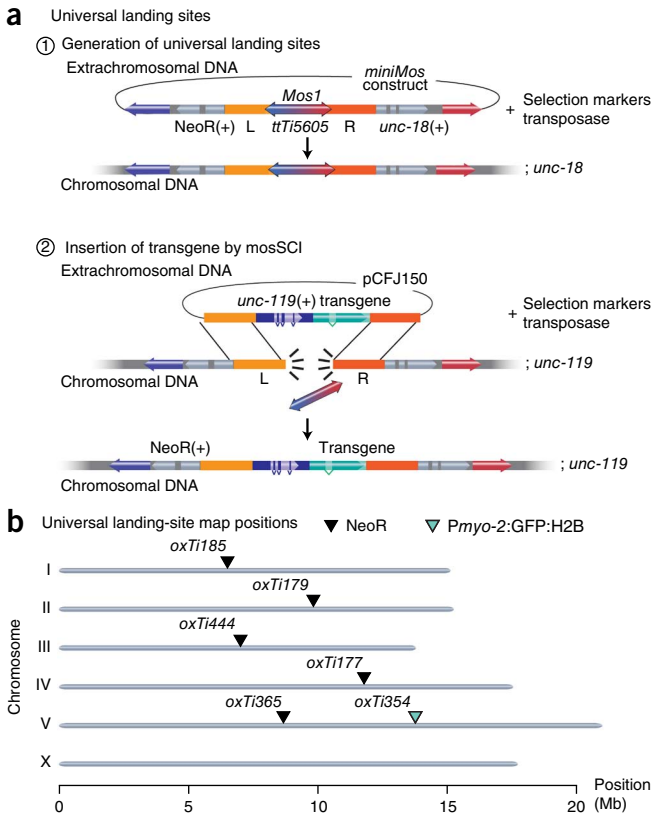


**Figure 3** | Using *miniMos* to generate universal mosSCI insertion sites. (**a**) Schematic of the method to generate universal mosSCI insertion sites. Step (1): insert *miniMos* with the *ttTi5605* genomic region (including the native *Mos1* element) into *unc-18* mutants. Cross inserts to *unc-119*. Step (2): inject the pCFJ150-based targeting vector to insert the transgene by mosSCI. All insertions were verified as functional, single-copy insertions. L, left, and R, right, recombination region from the *ttTi5605* mosSCI insertion site. (**b**) Genomic location of universal mosSCI insertion sites with verified germline expression. NeoR, neomycin-resistance gene.

screen for mutant phenotypes, we noted that several of the P*eft-3*:tdTomato:H2B insertions were inserted into introns and exons of genes with obvious phenotypes: *unc-13* I, *unc-22* IV and *him-4* X. All three insertions showed the phenotypes expected from loss-of-function alleles.

To test whether expression of insertions was affected by neighboring promoters, we generated strains with promoters driving GFP expression in pharyngeal muscles (P*myo-2*, $n = 3$) and body-wall muscle (P*unc-54*, $n = 3$). In this relatively small sample, we were unable to detect misexpression in other tissues (**Supplementary Fig. 5**). The insertion frequency and fidelity of insertions is robust enough that *miniMos* transposition could be a convenient alternative to extrachromosomal arrays in cases in which the unstable and multicopy nature of arrays is undesirable (**Supplementary Note** and **Supplementary Fig. 6**).

### *Mos1* can transpose with fosmids and *lacO* repeats

To determine the maximum cargo capacity of recombinant *Mos1* elements, we generated *Mos1*-based fosmids (Mosmids) by recombineering[36]. We inserted a cassette with a 1-kb recombinant *Mos1* element and *Cbr-unc-119*(+) into the backbone of several fosmids with GFP-tagged genes (**Fig. 2**). We injected five different Mosmids into *unc-119* animals and obtained stable integrated lines at $P_0$ frequencies ranging from 2% to 14% (5% ± 2%; mean ± s.e.m.) of all constructs. The drop-in insertion frequency was likely caused by two effects: larger cargo may inhibit transposition, and Mosmid injections only inefficiently form extrachromosomal arrays. Inserted Mosmids expressed EGFP in the expected tissues, including the germ line (**Fig. 2b**).

From one Mosmid (*air-2*:EGFP) we obtained 18 independent insertions that were all fluorescent, which suggests that Mosmid insertions were generally intact. We verified the integrity of the inserted fosmids by comparative genome hybridization (CGH); this method can detect deletions, insertions and even single-base-pair mutations with high sensitivity[37,38] (**Fig. 2c** and **Supplementary Fig. 7**). In the four lines generated from a tagged *cnd-1* gene, either a single, fully intact copy or two full copies (into a single location) of the Mosmid were inserted. We observed similar full-length insertions by CGH on lines from *gpb-1*, *his-55* and *air-2* inserts (**Supplementary Fig. 7**).

*lacO* repeats can be used to visualize chromosome position when they are bound to a fluorescently tagged LacI repressor[39]. We tested whether a recombinant *Mos1* element could insert a large repetitive transgene containing 256× *lacO* repeats and selection markers. We generated 20 independent insertions (**Supplementary Fig. 8**). These strains showed two distinct fluorescent dots in embryos when crossed into a line expressing LacI:GFP, corresponding to the two homologous chromosomes containing the *lacO* repeats (P. Meister, University of Bern, personal communication).

These experiments showed that the *miniMos* element is compatible with a wide variety of transgenic cargo and selection markers. We have generated a set of 16 standardized *miniMos* cloning vectors to facilitate use of the technique (**Supplementary Fig. 9**).

### A set of universal mosSCI insertion sites

The ΦC31 recombinase has been used in flies to develop universal insertion sites that are compatible with a single targeting vector[4,40]. We unsuccessfully attempted to adapt the ΦC31 system for *C. elegans* (M.S. and C.F.-J., unpublished observations). As an alternative, we developed a *miniMos* system that achieves the same goal. We generated a *miniMos* element containing the *ttTi5605* mosSCI site and flanked it with two selection markers, *unc-18* and either NeoR or P*myo-2*:GFP:H2B (**Fig. 3**). The embedded *ttTi5605* Mos element within the *miniMos* transposon can be used as a landing site for single-copy insertion using mosSCI[12] and is compatible with previously published targeting vectors (pCFJ150 or pCFJ350) (**Fig. 3**). Furthermore, mosSCI insertions can be followed in crosses by the adjacent selection marker (NeoR or P*myo-2*:GFP:H2B). We generated a set of validated single-copy, full-length mosSCI universal insertion sites that were permissive for germline expression (**Fig. 3**). Additionally, we targeted the insertion of a universal landing site into the *ttTi25545 Mos1* site at the center of chromosome III by mosSCI because no insertion site on chromosome III was compatible with germline expression (data not shown). All universal landing sites were validated: we could generate single-copy inserts at frequencies similar to those for insertions into the native *ttTi5605* site, and a P*dpy-30*:GFP:H2B transgene was expressed in the germ line (**Supplementary Table 1**).

## DISCUSSION

Random insertion of transgenes with the *miniMos* element has several advantages relative to biolistic transformation[8]. First, the exact insertion site can be determined by PCR. Knowledge of the exact insertion site ensures that mutations caused by *miniMos* insertion, or effects on expression of the transgene by the genomic environment, can be assessed. Second, a single intact copy of the transgene with well-defined end points in the genome is inserted. Third, the *miniMos* element can insert intact fosmids[41] and is active in other species and natural *C. elegans* isolates[42]. Finally, the insertion frequency of the *miniMos* element is high enough that several insertions are frequently generated from a single injection. Redundant inserts improve the chance of identifying insertions that do not disrupt endogenous genes and that are appropriately expressed.

We imagine *miniMos* transgenesis will mostly be used to insert single copies of transgenes, but there are at least four additional uses for the *miniMos* resources described here. (i) The set of dominant chromosome balancers is composed of 158 inserts that express red or green fluorescent proteins in somatic nuclei spaced about every 2–5 map units (**Supplementary Fig. 4**). These balancers can be used to generate strains with complicated genotypes. (ii) We generated two mapping strains that contain three distinguishable fluorescent markers that cover all six chromosomes in high incidence of male (*him*) mutant backgrounds. These strains are useful for mapping new mutations to chromosomes. (iii) The *lacO* insertions mark 20 different genomic sites and can be used to locate chromosome positions in the nucleus: for example, during meiosis or differentiation[43]. (iv) We generated a set of universal mosSCI insertion sites that are compatible with a single targeting vector. These strains can be used to insert single-copy transgenes at multiple positions in the genome.

In the future, two compelling uses for *miniMos* will be to probe the genome on a global scale for chromatin effects and to determine expression patterns using gene-trap constructs. First, the preliminary experiments with the composite *Mos* inserts demonstrate that transgene expression in both the soma and germ line of *C. elegans* is position dependent, with high degrees of silencing on the X chromosome and on autosomal arms. For example, almost all of the nonfluorescent P*pie-1*:GFP insertions were inserted into the X chromosome, which is inactivated in the germ line[23], or into autosomal arms containing a high incidence of repressive histone marks[22]. Second, *miniMos* constructs can be used to generate enhancer-trap and gene-trap constructs. For determining the expression pattern of a single gene, it will be much more efficient to specifically target the gene with the CRISPR-Cas9 system[15,16,44]. But for determining the expression patterns of all genes, random insertions with *miniMos* will be preferable, as has been done in *Drosophila* using P elements[2]. The *miniMos* element could be combined with the Q system[45] to generate strong, inducible driver lines for most tissues. In particular, it may be possible to identify promoters or enhancers that target expression individually to many of the 302 neurons of the adult nervous system.

Protocols, annotated plasmid sequences and a searchable list of strains are available at the Wormbuilder web page (http://www.wormbuilder.org/).

## METHODS

Methods and any associated references are available in the online version of the paper.

*Note: Any Supplementary Information and Source Data files are available in the online version of the paper.*


### ACKNOWLEDGMENTS

We thank B. Waterston (University of Washington), A. Sapir and P. Sternberg (California Institute of Technology), and the NemaGENETAG consortium for strains; B. Meyer (UC Berkeley) and P. Meister (University of Bern) for validating lacO insertions; the J. Chin (MRC, University of Cambridge), D. Dupuy (University of Bordeaux), B. Lehner (EMBL-CRG, Systems Biology Unit, Barcelona) and G. Seydoux (John Hopkins University) labs for plasmids; M. Maduro (UC Riverside) for improving mosSCI insertion frequency; and K. Hoe for expert technical assistance. Some strains were provided by the *Caenorhabditis* Genetics Center (CGC), which is funded by US National Institutes of Health (NIH) Office of Research Infrastructure Programs (P40 OD010440). This work was supported by the Carlsberg Foundation (C.F.-J.), NIH grant 1R01GM095817 (E.M.J.), US National Science Foundation grant NSF IOS-0920069 (E.M.J.) and the Howard Hughes Medical Institute (E.M.J.). The Mosmid engineering work was supported by the Max Planck Society (MPG) Initiative "BAC TransgeneOmics" and the NIH ModENCODE project. Work in the laboratory of D.G.M. was supported by the Canadian Institute for Health Research. Work in the laboratory of D.G.M. was supported by the Canadian Institute for Health Research and the Canadian Institute for Advanced Research.


### AUTHOR CONTRIBUTIONS

C.F.-J. designed experiments under the supervision of E.M.J. and M.W.D. C.F.-J., M.S., A.P., J.T., M.L. and S.F. performed the research. C.F.-J. performed molecular biology, injections, imaging and genetics; M.L. generated mapping strains; M.S. and A.P. performed fosmid recombineering; and J.T., S.F. and D.G.M. performed comparative genome hybridization. C.F.-J. and E.M.J. wrote the paper with input from all coauthors.

### COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.




1. Rubin, G.M. & Spradling, A.C. Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**, 348–353 (1982).
2. Brand, A.H. & Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401–415 (1993).
3. Spradling, A.C. *et al.* Gene disruptions using P transposable elements: an integral component of the *Drosophila* genome project. *Proc. Natl. Acad. Sci. USA* **92**, 10824–10830 (1995).
4. Venken, K.J.T., He, Y., Hoskins, R.A. & Bellen, H.J. P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in *D. melanogaster*. *Science* **314**, 1747–1751 (2006).
5. Wallrath, L.L. & Elgin, S.C. Position effect variegation in *Drosophila* is associated with an altered chromatin structure. *Genes Dev.* **9**, 1263–1277 (1995).
6. Ivics, Z. *et al.* Transposon-mediated genome manipulation in vertebrates. *Nat. Methods* **6**, 415–422 (2009).
7. Mello, C.C., Kramer, J.M., Stinchcomb, D. & Ambros, V. Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. *EMBO J.* **10**, 3959–3970 (1991).
8. Praitis, V., Casey, E., Collar, D. & Austin, J. Creation of low-copy integrated transgenic lines in *Caenorhabditis elegans*. *Genetics* **157**, 1217–1226 (2001).
9. Bessereau, J.-L. *et al.* Mobilization of a *Drosophila* transposon in the *Caenorhabditis elegans* germ line. *Nature* **413**, 70–74 (2001).
10. Robert, V. & Bessereau, J.-L. Targeted engineering of the *Caenorhabditis elegans* genome following *Mos1*-triggered chromosomal breaks. *EMBO J.* **26**, 170–183 (2007).
11. Frøkjær-Jensen, C., Davis, M.W., Ailion, M. & Jorgensen, E.M. Improved *Mos1*-mediated transgenesis in *C. elegans*. *Nat. Methods* **9**, 117–118 (2012).
12. Frøkjaer-Jensen, C. *et al.* Single-copy insertion of transgenes in *Caenorhabditis elegans*. *Nat. Genet.* **40**, 1375–1383 (2008).

13. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
14. Chen, C., Fenk, L.A. & de Bono, M. Efficient genome editing in *Caenorhabditis elegans* by CRISPR-targeted homologous recombination. *Nucleic Acids Res.* **41**, e193 (2013).
15. Dickinson, D.J., Ward, J.D., Reiner, D.J. & Goldstein, B. Engineering the *Caenorhabditis elegans* genome using Cas9-triggered homologous recombination. *Nat. Methods* **10**, 1028–1034 (2013).
16. Friedland, A.E. *et al.* Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. *Nat. Methods* **10**, 741–743 (2013).
17. Lohe, A.R. & Hartl, D.L. Efficient mobilization of *mariner in vivo* requires multiple internal sequences. *Genetics* **160**, 519–526 (2002).
18. Lozovsky, E.R., Nurminsky, D., Wimmer, E.A. & Hartl, D.L. Unexpected stability of *mariner* transgenes in *Drosophila*. *Genetics* **160**, 527–535 (2002).
19. Horn, C. & Wimmer, E.A. A versatile vector set for animal transgenesis. *Dev. Genes Evol.* **210**, 630–637 (2000).
20. Casteret, S. *et al.* Physical properties of DNA components affecting the transposition efficiency of the *mariner Mos1* element. *Mol. Genet. Genomics* **282**, 531–546 (2009).
21. Boulin, T. & Bessereau, J.-L. *Mos1*-mediated insertional mutagenesis in *Caenorhabditis elegans*. *Nat. Protoc.* **2**, 1276–1287 (2007).
22. Liu, T. *et al.* Broad chromosomal domains of histone modification patterns in *C. elegans*. *Genome Res.* **21**, 227–236 (2011).
23. Meyer, B.J. Targeting X chromosomes for repression. *Curr. Opin. Genet. Dev.* **20**, 179–189 (2010).
24. Seth, M. *et al.* The *C. elegans* CSR-1 argonaute pathway counteracts epigenetic silencing to promote germline gene expression. *Dev. Cell* **27**, 656–663 (2013).
25. Shirayama, M. *et al.* piRNAs initiate an epigenetic memory of nonself RNA in the *C. elegans* germline. *Cell* **150**, 65–77 (2012).
26. Wedeles, C.J., Wu, M.Z. & Claycomb, J.M. Protection of germline gene expression by the *C. elegans* argonaute CSR-1. *Dev. Cell* **27**, 664–671 (2013).
27. Vallin, E. *et al.* A genome-wide collection of *Mos1* transposon insertion mutants for the *C. elegans* research community. *PLoS ONE* **7**, e30482 (2012).
28. Lohe, A.R. & Hartl, D.L. Efficient mobilization of *mariner in vivo* requires multiple internal sequences. *Genetics* **160**, 519–526 (2002).
29. Maduro, M. & Pilgrim, D. Identification and cloning of *unc-119*, a gene expressed in the *Caenorhabditis elegans* nervous system. *Genetics* **141**, 977–988 (1995).
30. Giordano-Santini, R. *et al.* An antibiotic selection marker for nematode transgenesis. *Nat. Methods* **7**, 721–723 (2010).
31. Semple, J.I., Garcia-Verdugo, R. & Lehner, B. Rapid selection of transgenic *C. elegans* using antibiotic resistance. *Nat. Methods* **7**, 725–727 (2010).
32. Radman, I., Greiss, S. & Chin, J.W. Efficient and rapid *C. elegans* transgenesis by bombardment and hygromycin B selection. *PLoS ONE* **8**, e76019 (2013).
33. Kelly, W.G., Xu, S., Montgomery, M.K. & Fire, A. Distinct requirements for somatic and germline expression of a generally expressed *Caenorhabditis elegans* gene. *Genetics* **146**, 227–238 (1997).
34. Andersen, E.C. *et al.* Chromosome-scale selective sweeps shape *Caenorhabditis elegans* genomic diversity. *Nat. Genet.* **44**, 285–290 (2012).
35. Zhao, Z. *et al.* New tools for investigating the comparative biology of *Caenorhabditis briggsae* and *C. elegans*. *Genetics* **184**, 853–863 (2010).
36. Sarov, M. *et al.* A recombineering pipeline for functional genomics applied to *Caenorhabditis elegans*. *Nat. Methods* **3**, 839–844 (2006).
37. Maydan, J.S. *et al.* Efficient high-resolution deletion discovery in *Caenorhabditis elegans* by array comparative genomic hybridization. *Genome Res.* **17**, 337–347 (2007).
38. Maydan, J.S., Okada, H.M., Flibotte, S., Edgley, M.L. & Moerman, D.G. *De novo* identification of single nucleotide mutations in *Caenorhabditis elegans* using array comparative genomic hybridization. *Genetics* **181**, 1673–1677 (2009).
39. Robinett, C.C. *et al.* *In vivo* localization of DNA sequences and visualization of large-scale chromatin organization using Lac operator/repressor recognition. *J. Cell Biol.* **135**, 1685–1700 (1996).
40. Groth, A.C., Fish, M., Nusse, R. & Calos, M.P. Construction of transgenic *Drosophila* by using the site-specific integrase from phage φC31. *Genetics* **166**, 1775–1782 (2004).
41. Sarov, M. *et al.* A genome-scale resource for *in vivo* tag-based protein function exploration in *C. elegans*. *Cell* **150**, 855–866 (2012).
42. Semple, J.I., Biondini, L. & Lehner, B. Generating transgenic nematodes by bombardment and antibiotic selection. *Nat. Methods* **9**, 118–119 2012).
43. Meister, P., Towbin, B.D., Pike, B.L., Ponti, A. & Gasser, S.M. The spatial dynamics of tissue-specific promoters during *C. elegans* development. *Genes Dev.* **24**, 766–782 (2010).
44. Chen, C., Fenk, L.A. & de Bono, M. Efficient genome editing in *Caenorhabditis elegans* by CRISPR-targeted homologous recombination. *Nucleic Acids Res.* **41**, e193 (2013).
45. Wei, X., Potter, C.J., Luo, L. & Shen, K. Controlling gene expression with the Q repressible binary expression system in *Caenorhabditis elegans*. *Nat. Methods* **9**, 391–395 (2012).

## ONLINE METHODS

**Reagents.** Please see the web page http://www.wormbuilder.org/ for annotated plasmid sequences, protocols and a searchable lists of strains. Plasmids are available from Addgene as a single kit (#1000000031; https://www.addgene.org/minimos/) or as individual plasmids. Strains were maintained using standard methods[46]. Temperature-sensitive strains *lin-5* and *spd-1* were grown at 15 °C. All other strains were grown at room temperature on OP50 or HB101 bacteria. Fluorescent balancer strains, including the two mapping strains, have been deposited with the *Caenorhabditis* Genetics Center (CGC).

**Molecular biology.** Plasmids were designed with ApE (A plasmid Editor, M.W. Davis), which is freely available at http://www.biology.utah.edu/jorgensen/wayned/ape/.

All plasmids were generated by standard molecular techniques, including isothermal assembly[47] and three-fragment Gateway cloning (Life Technologies). PCR amplification was performed using a high-quality DNA polymerase, Phusion (New England BioLabs).

Please see **Supplementary Table 1** for GenBank-formatted plasmid sequences of all plasmids used in this study.

**Reproducibility.** All injections were performed at least in duplicate and usually in triplicate on different days. Only injections with DNA isolated by the same preparation method were compared. The number of injections and the sample size were selected to reach statistical significance in tests that correct for multiple comparisons. Overall, the reproducibility on different days was high. This is particular apparent in the experiment to identify the minimal *Mos1* element (*miniMos*), where all truncated constructs larger than the *miniMos* transposon show reproducible insertion frequencies (**Supplementary Fig. 1**).

*Exclusion criteria.* Plates that did not contain any transgenic $F_1$ progeny as determined by phenotypic rescue (*unc-119* injections) or the presence of fluorescent co-injection markers (antibiotic injections) were not counted toward the number of injected animals. This exclusion criteria excluded approximately 5–10% of all injected animals and served to reduce the variability caused by differences in injection needles between separate injections.

*Blinding and randomization.* No blinding or randomization was performed.

**Recombinant *Mos1* insertions.** *miniMos insertions*. Insertions were generated and mapped as described in detail in the **Supplementary Protocol**. In brief, injection strains were maintained on HB101 bacteria at 15–20 °C. An injection mix containing the *miniMos* transgene at 10–15 ng/µl, red fluorescent co-injection markers pGH8 at 10 ng/µl, pCFJ90 at 2.5 ng/µl and pCFJ104 at 10 ng/µl, a helper plasmid expressing the *Mos1* transposase pCFJ601 at 50 ng/µl and the negative, heat shock–inducible *peel-1* selection marker pMA122 at 10 ng/µl. The remaining volume was made up of milliQ purified water. Injected worms were placed at room temperature for 1–2 h, transferred to individual plates and incubated at 25 °C until starvation (approximately 1 week). For experiments aimed at quantifying insertion frequency, plates were screened for $F_1$ rescue 3 d after injection, and plates with no $F_1$ rescue were discarded. Once starved, plates were heat shocked for 2 h at 34 °C or for 1 h at 37 °C in an air incubator to kill animals with extrachromosomal arrays. All plates were screened for *miniMos* insertions the day after heat shock on a fluorescence microscope on the basis of rescue and the absence of red co-injection markers. Because of obvious visual differences (state of animals at 25 °C vs. 15 °C or the fluorescence of injected plasmids), the investigator was not systematically blinded to the injected constructs. A single animal from each plate containing insertions was picked for further analysis. The location of *miniMos* elements was determined by an inverse PCR protocol modified from Boulin and Bessereau[21] on genomic DNA isolated with the kits "ZR Tissue and Insect DNA miniprep" or "ZR-96 Genomic DNA Tissue miniprep" (Zymo Research). The DNA was digested with restriction enzymes (New England BioLabs) for 3 h to overnight, ligated with T4 ligase (Enzymatics) and PCR amplified twice with oligos that anneal in the *miniMos* transposon with Phusion DNA Polymerase. The PCR product was electrophoresed on a 1% agarose gel, and single bands were gel purified with the "Zymoclean Gel DNA Recovery Kit" (Zymo Research). The gel-purified product was Sanger sequenced at the University of Utah Sequencing Core.

We performed two or three independent injections for each set of conditions tested (for example, temperature or length of composite *miniMos* transposon) to minimize effects of a single bad injection needle. Generally, we observed very little variability between independent injections. Following advice from M. Maduro (UC Riverside), we determined that the largest source of variability was in the quality of injected DNA. We isolated DNA with Spin Miniprep (cat. no. 27106) and Plasmid Plus Midiprep (cat. no. 12943) kits from Qiagen and with a PureLink HQ Mini Plasmid kit from Invitrogen (cat. no. K2100-01). The higher-quality DNA kits (Qiagen Midi and Invitrogen Mini kits) resulted in a fourfold increase in $F_1$-rescued animals (20 vs. 5 rescued animals per injection) and a 50% (Qiagen Midi) to 100% (Invitrogen mini) increase in mosSCI insertion frequency (**Supplementary Fig. 6**). Although we have not tested the effect of DNA purity on *miniMos* insertion frequency, we generally recommend using DNA of higher purity for injection than what is isolated with the standard Qiagen Miniprep Kit. At the time of injections performed to quantify the insertion frequency of the *miniMos* transposon, we were not aware of the increased frequency resulting from higher DNA quality, and these injections were therefore all done with the Qiagen miniprep kit.

*Quantification of insertions per injection* (**Table 1**). We injected a mix of three different *miniMos* plasmids carrying P*eft-3*:GFP:H2B, P*eft-3*:mCherry or P*eft-3*:tdTomato:H2B with the *Cbr-unc-119*(+) selection together with the *Mos1* transposase and the negative PEEL-1 selection plasmid into *unc-119* mutant animals. We picked rescued animals in the $F_1$ generation to individual plates and allowed the animals on these plates to starve out at 25 °C. We heat-shocked plates with rescued $F_2$ or $F_3$ animals to kill animals with extrachromosomal arrays and screened for insertions the following day. We screened each plate containing an insertion for the presence of multiple different fluorescent patterns and picked a single animal from each plate for further analysis. We isolated genomic DNA and performed inverse PCR on all seven different insertions (*oxTi306*–*oxTi312*) that originated from injection into $P_0$ animal no. 5. All seven insertions mapped to different genomic locations.

*Universal insertion sites.* The universal insertion sites were generated by injection into *unc-18*(*md299*) animals following the protocol for *miniMos* insertions. The internal *Mos1* element depressed *miniMos* insertion frequency from approximately 60% to 12% ($n$ = 180) and resulted in a high frequency of complex insertions (56%, $n$ = 23). Strains with a putative insertion were tested for antibiotic resistance to G418 (NeoR). Genomic DNA was isolated from homozygous, G418-resistant strains and tested by PCR for the presence of the *ttTi5605 Mos1* element and the absence of backbone fragments from the cloning vector. Inverse PCR was performed on strains with intact universal insertion sites with oligos that specifically detect the *miniMos* element and not the wild-type (internal) *Mos1* element. The genomic location was determined by Sanger sequencing and verified by oligos designed for each individual insertion (**Supplementary Table 1**). Strains with universal insertion sites were outcrossed five times against an 11× outcrossed *unc-119*(*ed3*) strain, EG6207, derived from PS6038 (a kind gift from A. Sapir and P. Sternberg (Caltech)) by following neomycin resistance. We verified homozygosity of the universal insertion sites in the *unc-119* background after out-crossing by PCR. The ability to insert transgenes into all universal landing sites was verified by insertion of pCFJ150-derived constructs with P*eft-3*:GFP:H2B:*tbb-2* UTR, P*dpy-30*:GFP:H2B:*tbb-2* UTR or P*pie-1*:GFP:H2B:*pie-1* UTR transgenes.

In one case, *oxTi444*, a universal insertion site was generated by targeted insertion of the universal landing site into a preexisting mosSCI site, *ttTi25545*. In this case, the *miniMos* element was exchanged for left and right homology regions adjacent to *ttTi25545* and inserted by the standard mosSCI protocol[11].

*Antibiotic selection protocol.* We used antibiotic selection protocols modified from Giordano-Santini *et al.*[30], Semple *et al.*[31] and Radman *et al.*[32]. For G418 selection, we made a 25 mg/ml (Gold Biotechnology) solution in water and filter-sterilized the solution with a 0.2-µm filter. For puromycin selection we purchased a 10 mg/ml solution (InvivoGen) and added 0.1% Triton X-100 (Sigma). For hygromycin B we made a 20 mg/ml (Gold Biotechnology) solution in water and filter-sterilized the solution with a 0.2-µm filter. For use in antibiotic selection, 500 µl of the stock solutions were added directly to plates containing wild-type worms that had been injected 1 or 2 d before. Plates were allowed to dry with the lid off. Dry plates were returned to the 25 °C incubator, and worms were allowed to starve. The animals were heat shocked to remove those with extrachromosomal arrays and were screened for insertions the next day on the basis of survival on antibiotic plates, lack of fluorescent co-injection markers and fluorescence from the *miniMos* construct carrying P*eft-3*:GFP(NLS). At least ten animals from each antibiotic selection were propagated and homozygosed by fluorescence to verify true insertions. We note that the antibiotic selection markers are very convenient for injecting into healthier strains, such as wild-type animals, but suffer from the disadvantage that they are harder to homozygose, especially in the absence of a fluorescent insertion marker. In our hands, G418 and hygromycin B killed almost all nontransgenic animals within 2 d, whereas puromycin typically took 3–4 d to kill nontransgenic animals.

*Composite* Mos1 *remobilization.* To determine whether composite *Mos1* insertions can be remobilized from genomic locations, we generated a strain carrying an insertion (*oxTi51*; **Fig. 1b**) and a mutation in the *unc-18* gene. A rescuing template containing *unc-18*(+) was constructed so that a double-strand break generated by transposon excision would be repaired by homologous recombination and copy *unc-18*(+) into the excision site. From 48 injected animals we did not recover any targeted *unc-18*(+) insertions. This result is in agreement with similar experiments in *Drosophila*, where the insertion frequency was intact but genome mobilization was reduced by two orders of magnitude for modified transposons of the same family as *Mos*[18].

*Bioinformatic analysis of recombinant* Mos1 *insertions.* The locations of transposons were determined by inverse PCR. Genomic location was determined by identifying the junction between the transposon and genomic DNA. A BLAST search at http://www.wormbase.org/ against genome version WS190 (ce6) was used to determine the genomic position. Generally only uniquely identified insertions were used; however, some insertions that map to several position within a small genomic interval (~10 kb) were included in some figures.

**Comparative genome hybridization.** Genomic DNA from worms was isolated with the ZR Tissue & Insect DNA MiniPrep kit (Zymo Research) following the manufacturer's protocol. DNA labeling, sample hybridization, image acquisition and determination of fluorescence were all performed as previously described[37,38]. We used a 3× high-density (HD) chip divided into three whole-genome sections with 720,000 different oligos for all experiments. The chip design was based on our original whole-genome chip containing 385,000 different oligos. All microarrays were manufactured by Roche-NimbleGen with oligonucleotides synthesized at random positions on the arrays. The chip design name is 90420_Cele_RZ_CGH_HX3. Quantile normalization was performed on the intensity ratios for all experiments. Seven strains—EG7784 (*oxTi97*), EG7785 (*oxTi98*), EG7786 (*oxTi99*), EG7787 (*oxTi100*), EG6840(*oxTi109*), EG6731 (*oxTi114*) and EG6788 (*oxTi118*)—were tested against wild-type DNA. All strains showed a duplication of the full genomic region contained within the recombineered fosmid, except for the strain EG7787, which contains a dual insertion. PCR amplification from EG7787 showed the presence of backbone DNA, which is consistent with a duplicate insertion into the same genomic locus. For all analyzed Mosmid insertions, the end points of genomic duplications identified by CGH closely matched the ends of recombineered fosmids, and no second-site duplications were detected.

**Fosmid recombineering.** The fosmids were engineered essentially as in ref. 41, except for the fosmid backbone modification step, where the *Mos1* transposon (1,000 bp) with inverted repeats (IR) was added to the *Cbr-unc-119*-Nat cassette (on each side of the NatR marker). To make the fosmid host bacteria EPI300 (Epicentre) proficient for recombineering, we transformed the EPI300 cells with the pRedFlp4 plasmid, which allows for inducible expression of either the λ Red operon+RecA or the Flp recombinase. For gene tagging, a multipurpose tagging cassette that contains the flexible linker peptide TY1, GFP, *FRT*-flanked positive selection (NeoR), counterselection (rpsL) and the affinity tag 3xFlag was PCR amplified. The PCR used gene-specific primer extensions of 50 bp upstream and downstream of the insertion point that serve as homology arms for recombineering. Recombinants were selected for kanamycin resistance in liquid culture. The *rpsl/neo* selection-counterselection marker





doi:10.1038/nmeth.2889

was removed by Flp/FRT recombination. The homology arms targeting the *Cbr-unc-119*/IR NatR IR cassette to the fosmid backbone were the same for all fosmids and were included in the same plasmid (pCFJ496); this cassette was isolated by restriction digest from pCFJ496 and used for recombineering the fosmid containing a EGFP-tagged gene. Both the template for the multipurpose tagging cassette and the template for inserting the *Mos1* and *Cbr-unc-119* genes were cloned in plasmids with the R6K origin of replication, which is nonfunctional in the fosmid host strain, and removal of the plasmid is thus not required before recombineering. The fosmid modification cassette pCFJ496 is available from Addgene (plasmid #44488).

Mosmids generally integrate into the genome at lower frequencies than *miniMos* transposons that can be propagated as high-copy plasmids in bacteria. The lower insertion frequency is likely due to (i) lower transposition frequency of the *miniMos* element with larger cargo, (ii) decreased ability of fosmids to form extrachromosomal arrays owing to reduced homology and (iii) toxic sequences present on the fosmid. Some of the Mosmids that we tested were specifically chosen because integrated lines could not be generated by biolistic transformation despite repeated attempts and appear to be toxic (M.S., unpublished data). For example, we injected 48 and 60 *unc-119* animals with the *his-55*:EGFP and *his-56*:EGFP Mosmids, respectively. From these injections we did not recover a single rescued $F_1$ animal but were able to isolate one *his-55*:EGFP (2%) and two *his-56*:EGFP (3%) rescued insertion lines in the $F_2$ progeny. This suggests that these Mosmids are toxic at high copy number and that higher integration efficiencies may be achieved by titrating the Mosmid concentration. In support of this, we did not observe any toxicity from an *air-2*:EGFP Mosmid and recovered 18 independent insertions from 125 injected *unc-119* animals (14%).


46. Brenner, S. The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71–94 (1974).
47. Gibson, D.G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).